Iterative Roll Angle Estimation from Dense Disparity Map

Meghan Evans, Rui Fan, and Naim Dahnoun
Department of Electrical and Electronic Engineering, University of Bristol, Bristol, BS8 1UB, UK.
e-mail: meghan.c.evans@gmail.com

***Abstract*—The v-disparity map is predominantly used to estimate the parameters of the vertical profile of the road surface. Once the road surface is modelled, an object that lies away from it can be detected and determined as either an obstacle or a pothole. The accuracy of this estimation is largely affected by the clarity of the v-disparity map which can be vastly improved by eliminating the effect of a non-zero roll angle. With a rotation around the roll angle for the disparity map, a better v-disparity histogram can be provided. This paper presents a method for accurate roll angle estimation through analysis of the disparity and v-disparity maps. Since the quality of the v-disparity map is improved by rotating the disparity map by the estimated roll angle, this leads to improved road modelling. The more accurate the roll angle estimation, the larger this improvement is.**

I. Introduction

Robust and accurate object detection requires reliable road modelling. Imaging sensors such as digital cameras are usually used to obtain 3D images that can be used to monitor surrounding information. While some road fitting methods assume that the camera parameters are known [1] or the cameras are installed in such a way that their effects are negligible [2] [3], inaccurate estimation of camera parameters can compromise a system's reliability. Accurate estimation of roll angle is particularly important in systems that use the v-disparity map method [4]. This is due to the fact that with a larger roll angle, a wider range of disparities appear on each row of the disparity map [5]. This results in an unclear v-disparity map whose line parameters, and therefore road parameters, are harder to extract.

Changes in speed and road irregularities can cause the extrinsic rotations to change considerably [6]. To estimate the roll angle while in motion and without initial calibration, several methods whose inputs include sensors, lasers, and visual techniques are detailed in the literature. Visual techniques use either a monocular camera or a stereo camera, but the use of stereo vision is preferable as it provides 3D information which allows the roll angle to be determined with greater precision. Additionally, stereo vision allows for a much faster system [7].

The algorithm proposed in this paper furthers the work of Ozgunalp et al. [8], where the disparity map is estimated from a pair of images, and a plane is fitted to a small area of this disparity map to estimate the roll angle. The method used to determine the disparity map in the proposed algorithm is described by Fan et al. in [9]. In the original algorithm [8], a single patch was fitted per frame to a predetermined area. In this paper, a novel iterative process is proposed where multiple patches are fitted, with the final patch placed in the optimum position. This position is determined through the analysis of standard deviation across the entire map. Unlike the system described in [6], our algorithm operates only in the disparity domain. This is advantageous since no transformation in the Euclidean domain is required. To maintain a low complexity and to allow for a higher execution speed [10], integral images, otherwise known as summed-area tables, are used to calculate the standard deviation.

II. Algorithm Description

A. System Overview

Fig. 1 shows an overall block diagram of the proposed system. The algorithm proposed in [8] is implemented up to the point where the v-disparity map is obtained in Fig. 1 and the first estimation of the roll angle is the final output of the algorithm. The method proposed in this paper extends the algorithm in [8] using the subsequent steps in Fig. 1. Similar to the algorithm in [8], the shape of the initial patch is chosen as a square in the near-field as this area is less likely to include an object. However, single patch fitting [8] is not sufficient and it is necessary to find the optimum location for the fitted patch. This is because although using a patch fitted in the near-field area is likely to avoid objects, this area is very likely to include potholes which will, as with objects, not carry the same disparity as the road and so will cause inaccurate patch fitting.

To estimate the patch parameters, RANSAC (Random Sample Consensus) [11] is used to remove the effect of outliers. As with the method described in [12], a linear surface is then fitted to the points. Although the plane fitted at this point is linear, the road is ultimately modelled as a quadratic surface using the parameters determined from the v-disparity map. This is important as a linear surface would limit the detection of small objects or potholes whose detection could be affected by the curvature of the road. The parameters of the plane that is fitted to the patch are then used to determine the roll angle.

Once the roll angle has been determined, the image is rotated using the affine transformation, and a clearer v-disparity map is created from the new rotated image. From this v-disparity map, V_d, an expected disparity map, D_e, can be created with equal values of disparity across each row. The difference between this expected disparity map, D_e, and the

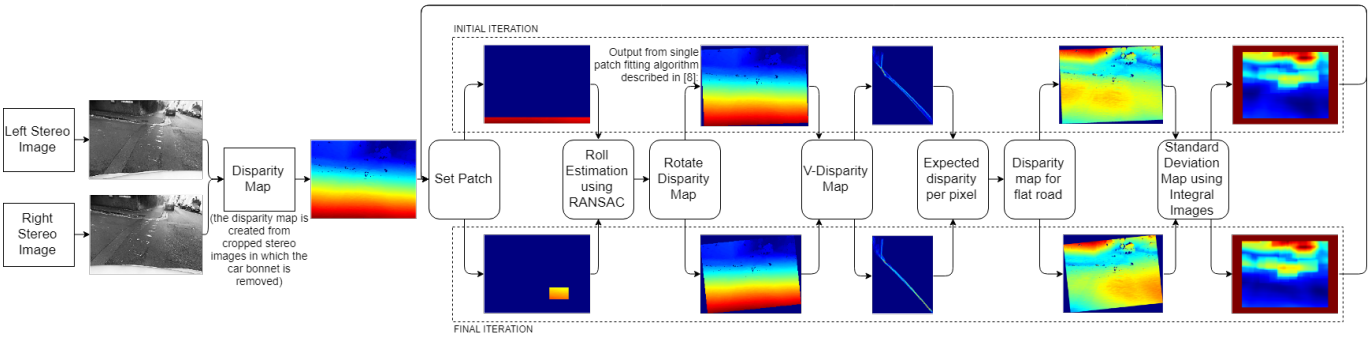

Fig. 1: Overall system block diagram.

rotated disparity map, D_r, is then used to create a standard deviation map, σ, from which the subsequent patch location is derived. The loop shown in Fig. 1 is followed through until the variation in roll angle between each iteration is reduced below a predetermined level. This level is dependent on the design requirements. For the tests carried out in the evaluation section of the paper, a value of 0.5° was used as this provided good results without compromising the execution speed. For systems where a higher accuracy is required, this threshold value can be reduced further at the expense of a lower execution speed. However, if the threshold is beyond a certain value, the condition may never be satisfied as the fluctuations in the roll angle will never reduce to zero.

B. Roll Angle Estimation Using RANSAC

The initial patch is positioned at the bottom of the disparity map in the near field. The column of the image in which the pixel is positioned is given by the *u*-axis. The *v*-axis corresponds to the row number and *d* represents the disparity. A linear plane is then fitted $(d = a_0 + a_1 u + a_2 v)$.
The roll angle, γ, can then be calculated as:

$$\gamma = \arctan \frac{-a_1}{a_2} \tag{1}$$

The plane is fitted using RANSAC according to the method described by Ozgunalp et al. in [8]. The sampled data points are discarded after use, and the process is repeated until the remaining outliers only make up of a small percentage of the total number of points. This threshold is again dependent on the design requirements where a compromise between accuracy and speed must be made. However, it is also dependent on the quality of the input disparity map. For low-quality disparity maps containing a high level of noise, the threshold number for outliers will need to be increased so that the outliers caused by noise do not prevent the condition from being reached. In this paper, the percentage of outliers was chosen as 0.1%.

The disparity map is then rotated by the roll angle, and the resulting map is used to create a v-disparity map. A quadratic polynomial curve is then fitted to the element with the highest value of each row in the v-disparity map using the least squares method.

$$d_e(u, v) = p_0 + p_1 v + p_2 v^2 \tag{2}$$

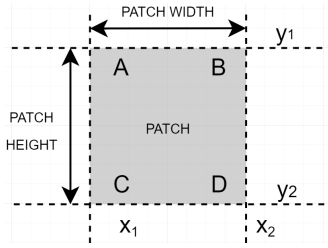

Fig. 2: Integral image, I.

where *v* is the row number of the pixel, d_e is the disparity of the pixel and p_0, p_1, p_2, are the parameters of the line given in the v-disparity map.

C. Integral Images for Standard Deviation Calculation

The disparities of the expected map, d_e, are subtracted from the rotated map, d_r, and a new map is created, $d(i,j) = (d_r(i,j) - d_e(i,j))$. This new map is used to determine the standard deviation as it removes the effect of the non-flat road surface. All points that lie on the estimated road surface carry a similar disparity. So, any increase in standard variation will be caused by either objects or potholes, and these need to be avoided when setting the patch location.

The method used to determine the integral images is as described by Fan et al. in [13]. The value held in each integral image at (x, y) is as follows:

$$S_d(x,y) = \sum_{i \le x, j \le y} d(i,j),\ S_{d^2}(x,y) = \sum_{i \le x, j \le y} (d(i,j))^2 \tag{3}$$

The standard deviation is computed as follows:

$$\sigma = \sqrt{\frac{1}{hw}(s_{d^2} - \frac{s_{d^2}}{hw})} \tag{4}$$

where $(h = (y_2 - y_1))and(w = (x_2 - x_1))$.

Using integral images, the complexity of the algorithm is reduced so that only four pixels need to be analysed, rather than all the pixels in the patch. This is because, as described in [14], the sum over the rectangle in Fig. 2 can now be calculated as shown in (5), where the references *A*, *B*, *C* and

D represent $I(x_1 - 1, y_1 - 1)$, $I(x_2, y_1 - 1)$, $I(x_1 - 1, y_2)$, and $I(x_2, y_2)$ respectively.

$$S = D + A - B - C \tag{5}$$

To determine the best position for the patch to be fitted, the sum of the standard deviation values within the patch needs to be calculated for the case of the patch being centred at all points in the disparity map. Without integral images, the value of all pixels in the patch area of the σ map need to be summed. However, by using the method described above only four pixels need to be analysed for each possible position. This significantly decreases the complexity of the algorithm. For an optimum patch width determined as in Fig. 3, the introduction of integral images means that only 0.02% of the pixel calculations are required.

D. V-Disparity for Road Model

The value in each pixel of the disparity map is inversely proportional to the distance from the camera. So, the v-disparity map, where each row is a histogram of the disparities that occur in the equivalent disparity map row, can be used to achieve what is effectively a side-on view of the road. This means that the v-disparity map can be used to model the road surface and detect obstacles or potholes. Points that lie on this road surface can be removed, leaving only the profile of the objects and potholes.

III. Experimental Evaluation

An artificial disparity map was used to measure the algorithm's ability to estimate an accurate roll angle. To create a disparity map that would be representative of real conditions and suitable for testing, a disparity map was first created with all rows containing an equal disparity. Then, Gaussian white noise was simulated and added to the image. To determine what effect an object lying in the area of the fitted patch would have on the roll angle estimation, a block of set disparity that was not equal to the disparity of the road was added to the map. Finally, the image was rotated by a known angle.

It was found that small to medium-sized objects that are most likely to be present in a real-world situation, did not affect the new algorithm since the patch was chosen to avoid them. Additionally, the single patch fitting method described in [8] is not affected by the introduction of objects unless they lay within the area of the patch fitted at the bottom of the disparity map. Therefore, since an added object would have no effect unless overlapping the patch at the bottom of the disparity map, only the case of overlap is considered. Fig. 3, Fig. 4, Fig. 5 and Fig. 6 compare the algorithm in [8] with the one proposed in this paper.

The performance of our new algorithm was largely dependent on the size of the patches that were fitted after the initial patch. This is shown in Fig. 3 where the performance of patches with a width below 201 pixels $((radius \times 2) + 1)$ is much worse than the performance of the single patch fitting. However, the performance of multiple patch fitting with a horizontal patch radius between 100 and 600 pixels is far

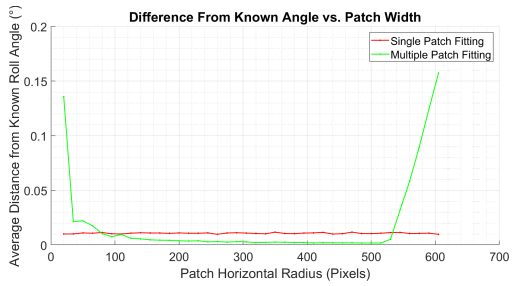

Fig. 3: Accuracy of roll angle estimation with varying patch width.

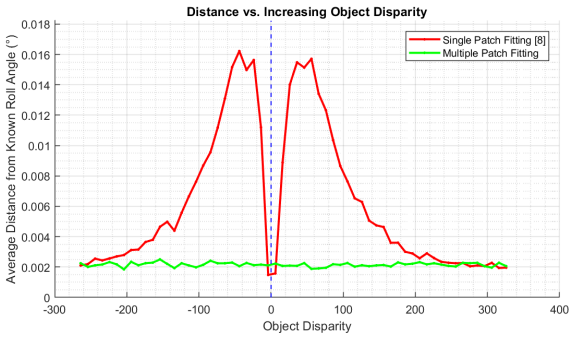

Fig. 4: Accuracy of roll angle estimation with varying difference between road disparity and object disparity.

superior, with the system performing at its best at a radius of 600 pixels. However, there is a tradeoff between speed and performance since a larger patch will take longer to analyse. The graphs in Fig. 4 and Fig. 5 were created using a horizontal patch radius of 400. Although the proposed algorithm does not run in real time, the authors believe that its execution speed can be boosted by exploiting the parallel computing architecture, such as multi-core processing [15].

Fig. 4 shows the average performance of the two algorithms over a varying object disparity. When an object is overlapping the initial patch, the single patch method's ability to estimate the roll angle accurately is compromised whereas this paper's algorithm performs well despite the change in disparity. The drop in the distance at a disparity difference of 0 is a result of the object disparity being very close to the disparity of the road in the patch. The red line drops below the green line at this point because the initial patch was larger than all subsequent patches as it spanned the width of the image. This initial patch is the only patch used in the case of single patch fitting.

A. Histogram

The single patch fitting and multiple patch fitting methods were then tested with a disparity map created from pictures taken of a real-world situation. The disparity map used in this test was developed from left and right camera images using the method described in [13] which delivered accurate disparity values. This original image is shown in Fig. 6a and contains a pothole in the near field which overlaps the initial patch area. This affects the fitting process for the single patch method.

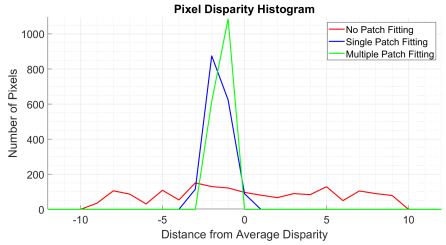

Fig. 5: V-disparity row histogram.

As with the previous experiment, the horizontal radius of the patches to be fitted after the initial patch was set at 400 pixels.

Fig. 6b and Fig. 6c show the images rotated using the single patch fitting and the multiple patch fitting respectively. While the difference between these two images is barely noticeable, the difference can be seen clearly in the corresponding v-disparity maps (Fig. 6e and Fig. 6f). Fig. 6d shows the v-disparity map for the unrotated image which is unclear with a wide range of disparities present on each row. Estimating the road parameters from this map would be challenging and inaccurate. When the image is rotated according to the angle determined with single patch fitting, the v-disparity map is improved. However, the final v-disparity map is even clearer and will, therefore, result in a more accurate road modelling.

The improvement becomes even more apparent when a single row of the v-disparity map is inspected and used to create a histogram, as shown in Fig. 5. In this case, the middle row of the image has been analysed, and the zero point in the axis represents the average of the disparities present in the row. The red line shows that the spread of disparities that appear in every row is large in the v-disparity map that is created from an unrotated image (Fig. 6d). There is also no clear maximum that can be used to fit a line to determine the parameters of the road. Both the single and multiple patch fitting improve on this. However, the multiple patch fitting outperforms the single patch fitting with a higher maximum and lower spread. The point on the v-disparity map row to which the quadratic equation should be fitted is very clear.

IV. Conclusion

The primary aim of the developed algorithm was to improve the quality of the v-disparity map to allow for more accurate estimation of the road parameters. The method proposed can estimate the roll angle more accurately and thus produce a clearer v-disparity map. However, it is an iterative process that analyses the standard deviation over the entire area of the disparity map and is therefore limited by the trade-off between roll angle precision and computation speed.

The algorithm can be improved for real-time applications. This would require an increase in speed and one way to achieve this would be to determine the next patch based on the patch chosen for the previous frame. When finding the position of the next patch, a search range could be established that would limit the area over which the standard deviation needs to be calculated.

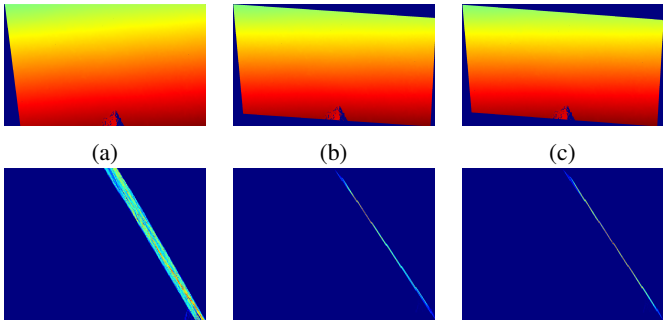

Fig. 6: Disparity maps with no fitting. (a), single patch fitting (b) and multiple patch fitting (c) with corresponding v-disparity maps (d), (e) and (f).

References

[1] U. Ozgunalp, X. Ai, and N. Dahnoun, "Stereo vision-based road estimation assisted by efficient planar patch calculation," *Signal, Image and Video Processing*, vol. 10, no. 6, pp. 1127–1134, 2016.

[2] M. Wu, S.-K. Lam, and T. Srikanthan, "Nonparametric technique based high-speed road surface detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 874–884, 2015.

[3] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, T. Graf, and R. Schmidt, "High accuracy stereovision approach for obstacle detection on non-planar roads," *Proc. IEEE INES*, pp. 211–216, 2004.

[4] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through" v-disparity" representation," in *Intelligent Vehicle Symposium, 2002. IEEE*, vol. 2. IEEE, 2002, pp. 646–651.

[5] N. Dahnoun, "Introduction to dsp," *Multicore DSP: From Algorithms to Real-time Implementation on the TMS320C66x SoC*, pp. 1–13.

[6] A. D. Sappa, F. Dornaika, D. Ponsa, D. Geronimo, and A. Lopez, "An efficient approach to onboard stereo vision system pose estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 3, pp. 476–490, 2008.

[7] S. Bota and S. Nedevschi, "Camera motion estimation using monocular and stereo-vision," in *Intelligent Computer Communication and Processing, 2008. ICCP 2008. 4th International Conference on*. IEEE, 2008, pp. 275–278.

[8] U. Ozgunalp, R. Fan, X. Ai, and N. Dahnoun, "Multiple lane detection algorithm based on novel dense vanishing point estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 621–632, 2017.

[9] R. Fan, X. Ai, and N. Dahnoun, "Road surface 3d reconstruction based on dense subpixel disparity map estimation," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3025–3035, 2018.

[10] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.

[11] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," in *Readings in computer vision*. Elsevier, 1987, pp. 726–740.

[12] Z. Zhang, X. Ai, C. Chan, and N. Dahnoun, "An efficient algorithm for pothole detection using stereo vision," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 564–568.

[13] R. Fan and N. Dahnoun, "Real-time implementation of stereo vision based on optimised normalised cross-correlation and propagated search range on a gpu," in *Imaging Systems and Techniques (IST), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–6.

[14] G. Facciolo, N. Limare, and E. Meinhardt-Llopis, "Integral images for block matching," *Image Processing On Line*, vol. 4, pp. 344–369, 2014.

[15] R. Fan, V. Prokhorov, and N. Dahnoun, "Faster-than-real-time linear lane detection implementation using soc dsp tms320c6678," in *Imaging Systems and Techniques (IST), 2016 IEEE International Conference on*. IEEE, 2016, pp. 306–311.